**Problem Set: 2 Regression Diagnostics**

1. The following table gives data on GNP and money supply (M1) in millions of dollars for Canada for the period 1970-1984.

|  |  |  |
| --- | --- | --- |
| Year | GNP (Y) | Money supply M1(X) |
| 1970 | 85,685 | 9,077 |
| 1971 | 94,450 | 10,178 |
| 1972 | 105,234 | 11,626 |
| 1973 | 123,560 | 13,320 |
| 1974 | 147,528 | 14,555 |
| 1975 | 165,343 | 16,566 |
| 1976 | 191,857 | 17,889 |
| 1977 | 210,189 | 19,381 |
| 1978 | 232,211 | 21,328 |
| 1979 | 264,279 | 22,823 |
| 1980 | 297,556 | 24,254 |
| 1981 | 339,797 | 25,379 |
| 1982 | 358,302 | 25,541 |
| 1983 | 390,340 | 28,137 |
| 1984 | 420,819 | 28,798 |

Use an appropriate model to the above data and comment on the results.

1. You are given the following data:

|  |  |  |  |
| --- | --- | --- | --- |
| Y | X | Y | X |
| 86 | 3 | 62 | 35 |
| 79 | 7 | 52 | 45 |
| 76 | 12 | 51 | 55 |
| 69 | 17 | 51 | 70 |
| 65 | 25 | 48 | 120 |

Fit the following model to the above data and obtain the usual regression statistics

.

1. The following table gives the data on stature (x) and sitting height (y) both in cm of each of 15 people of a certain Indian caste.

|  |  |
| --- | --- |
| **x** | **y** |
| 172.8  166.0  199.9  164.4  168.8  165.2  170.0  163.5  169.4  167.8  160.3  164.4  161.2  164.0  159.1 | 100.9  83.6  81.3  85.4  83.9  81.1  84.9  81.1  84.9  82.7  78.1  85.4  81.7  80.9  79.6 |

1. Draw the scatter plot of y versus x. Do you find anything unusual in the scatter plot?
2. Fit a linear regression equation of y on x and comment on the fit from the value of r2.
3. Compute the standardized residuals and plot those against x. Interpret the graph.
4. Indicate if there is any outlier in the data?
5. Compute the leverages and identify if there is any potential leverage point.
6. Compute Cook’s distance and indicate if there is any influential point.